ERDC/EL SR-18-1

Environmental Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

ERDC
INNOVATIVE SOLUTIONS for a safer, better world

Dredging Operations and Environmental Research (DOER) Program

Weight-of-Evidence Concepts: Introduction and Application to Sediment Management

Matthew E. Bates, Olivia C. Massey, and Matthew D. Wood

March 2018

Approved for public release; distribution is unlimited.

The U.S. Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at http://acwc.sdp.sirsi.net/client/default.

Cover artwork incorporates image "Information Overload" (https://commons.wikimedia.org/wiki/File:InformationOverload.jpg) by Eilidhm-cauley, used under CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0/deed.en) / Edges softened, and image "Direction Way Clouds Seat Belts" (http://maxpixel.freegreatpicture.com/Direction-Way-Clouds-Seat-Belts-427984), used under CC0 1.0

Dredging Operations and Environmental
Research (DOER) Program

ERDC/EL SR-18-1
March 2018

Weight-of-Evidence Concepts: Introduction and Application to Sediment Management

Matthew E. Bates and Matthew D. Wood

Environmental Laboratory
U.S. Army Engineer Research and Development Center
696 Virginia Road
Concord, MA 01742

Olivia Massey

Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

Final report

Approved for public release; distribution is unlimited.

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

Under Project 406028, "A Weight-of-Evidence Process for Environmental Assessments"

Abstract

The U.S. Army Corps of Engineers (USACE) has a variety of information available to support Civil Works project planning and operations. While this information must be interpreted to inform conclusions, clear guidance is often not available to describe how best to compare and integrate different types of information. This special report introduces a Weight-of-Evidence (WOE) approach that USACE staff can use to interpret the many Lines of Evidence (LOEs) available in the information based on the conclusions that they support and how much weight they each should have in the decision, helping to bridge the data-to-decisions gap. A case study is presented applying WOE to evaluate potential sediment placement sites for dredged material from New Haven Harbor, CT, the busiest port on Long Island Sound and one of the busiest ports in New England. The analysis demonstrates how diverse pieces of information can be evaluated based on their quality, resolution, and relevance and combined in the context of the problem at hand to aid robust and transparent USACE Civil Works decision-making.

DISCLAIMER: The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products. All product names and trademarks cited are the property of their respective owners. The findings of this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the Engineer Research and Development Center (ERDC) under the Dredging Operations and Environmental Research (DOER) Program; Project 406028, “A Weight-of-Evidence Process for Environmental Assessments.” The Program Manager was Dr. Todd S. Bridges (CEERD-EMD).

The work was performed by the Environmental Risk Branch (CEERD-EPR) of the Environmental Processes and Engineering Division (CEERD-EP), U.S. Army Engineer Research and Development Center, Environmental Laboratory (ERDC-EL). At the time of publication, Bill Nelson was Chief, CEERD-EPR; Warren P. Lorentz was Chief, CEERD-EP; and Dr. Todd S. Bridges, CEERD-EMD, was the Senior Research Scientist for Environmental Science. The Deputy Director of ERDC-EL was Dr. Jack Davis and the Director was Dr. Beth Fleming.

The scoring and analysis used in the case study represent the authors’ judgments and may not reflect the official views of the New England District (NAE) of the U.S. Army Corps of Engineers (USACE). The authors thank Craig Martin for sharing data on New Haven Harbor, Margaret Kurth for help summarizing prior applications, and Cate Fox-Lent, Craig Martin, and Dr. Burton Suedel for feedback on the text.

The Commander of ERDC was COL Bryan S. Green and the Director was Dr. David W. Pittman.

1 Introduction

1.1 Background

The U.S. Army Corps of Engineers (USACE) has a variety of information available to support Civil Works project planning and operations. Clear guidance, however, is often not available describing how to best compare and integrate different types of information to support decision making. For example, the jointly developed doctrine by the U.S. Environmental Protection Agency (USEPA) and USACE in the ocean and inland testing manuals (1991, 1998) provides guidance on how to test for evidence of contamination in dredged sediment, but is less explicit about how to synthesize, aggregate, trade off, and make sense of disparate testing results. This can be problematic when various lines of evidence come from sources with different levels of trust or suggest different conclusions regarding the most reasonable action. The types of judgments needed in these cases can be supported through transparent and quantitative *Weight-of-Evidence* (WOE) approaches that explicitly weight relevant properties of the data and data sources to provide context for synthesizing heterogeneous information and forming conclusions (Linkov et al. 2011).

1.2 Objective

Evidence available for use in decision making can vary in terms of its quality, relevance, and resolution, etc. These factors affect how strongly or weakly we believe the claims the evidence makes and provide important considerations for interpreting evidence in context and guiding data aggregation. The goal of this technical note is to illustrate how the different properties of data (metadata) can be leveraged to better integrate and interpret evidence in context and draw conclusions across differing lines of evidence.

1.3 Approach

The authors introduce the concept of Weight of Evidence (WOE) as an approach for synthesizing information and demonstrate its application through a case study on dredged material management for New Haven Harbor, CT.

In this case study, lines of evidence that vary in relevance, quality, and resolution are considered to evaluate potentially suitable alternatives for sediment placement. The example shows how application of the WOE methodology can help bridge the data-to-decision gap.

2 What Is Evidence?

When we refer to evidence in the context of USACE Civil Works projects, we mean one or more pieces of information that may change a decision maker's preference for one or more project alternatives under consideration. In the case of dredged material management, this could be a binary yes or no decision to advance to the next tier of testing or to use or not use a certain placement area; it could also be a more continuous decision, such as the density and number of sediment samples to take in a proposed dredge channel or the proportion of dredged sediments to place in an area. Each piece of evidence should help the decision maker update his or her belief about the suitability of one or more decision alternatives.

We can think about these pieces of information being organized into *Lines of Evidence* (LOEs), where each line corresponds to a distinct type of information. For instance, if a manager wants to determine the suitability of a placement site for use with a future dredging project, one LOE might relate to historic use of that placement site; other LOEs could cover past bulk chemistry testing records at the dredge and disposal sites, and so on. Each of these types of information tells us something that is roughly independent of the information conveyed by other LOEs.

Once evidence is categorized, there are two ways that it can influence decision making – as a result of the conclusions that can be directly drawn from it either supporting or opposing some hypothesis ("what the evidence says") and as a result of metadata about the evidence that suggest how much/little we should let it influence our overall conclusions ("how strongly the evidence says it" or "how much we believe it").

2.1 What is evidence metadata?

Evidence metadata describe the properties of each LOE and can be useful for interpreting evidence in context. Here, we describe three types of metadata in detail for use in a case study: evidence relevance, evidence quality, and evidence resolution. Other types of metadata (briefly mentioned below but not described in detail) can be included in different analyses, depending on the scope and data availability.

Relevance: Relevance is the first property we describe for determining whether and how much to allow a specific piece of evidence to influence a

decision. An LOE is more relevant to a decision if its inclusion will lead to a greater change in preference among the decision alternatives than would be found from receiving other information. For example, records of industrial discharges and maritime tanker accidents along a dredging channel may be more relevant when determining whether or not to do bulk chemistry sampling than the origin and contents of routine shipping traffic. The relevance of each LOE can be assessed along a continuum. For example, in comparing three types of evidence for an LOE related to bulk chemistry testing at an inland dredge site, recent sampling results from the site may be most relevant, followed next by historic site usage information, and followed lastly by historic sampling results from a downstream location. While joint EPA/USACE guidance (1991, 1998) assists managers in identifying LOEs that are likely to be relevant, additional transparent judgments need to be made to identify some LOEs as more relevant than others for supporting the decision at hand.

Quality: Evidence quality is also important for deciding how much weight specific information should carry in a decision. Evidence quality refers to the extent to which the methods used to produce the evidence are detailed, defensible, state-of-the-art, and unlikely to produce error. High-quality LOEs are often from objective and reliable sources, while low-quality LOEs may be less informative because they are accompanied by significant skepticism or uncertainty. For example, data logs from an industrial effluent sensor managed by a local regulatory agency might constitute high-quality evidence because we judge the data to have been gathered through consistent and state-of-the-art means (sensors); we trust that the agency would have had good access to place the sensor in a location necessary to make reliable measurements reflecting actual discharges; and we think the agency is likely consistent in how they report discharge measurements over time. Lower-quality evidence, on the other hand, might come from a community group’s downstream observations about the number and type of discharges, because we expect greater variation between different observing members using a means of estimation that is more susceptible to human error (visual readings).

Resolution: Another important property that can help inform evidence’s usefulness is its resolution. Resolution refers to the density and/or quantity of observations within the evidence. For example, high-resolution data might come from a moored flow rate sensor in a shipping channel that provides continuous output. By comparison, low-resolution data

might come from a portable device that a worker uses to make flow rate measurements once or twice a month. The number of significant figures in a dataset (if used responsibly) may similarly reflect resolution. While high-resolution data may also tend to be of higher quality, this is not always the case. Going back to the flow rate example, a poorly designed flow meter that provides continuous readings but is prone to providing incorrect ones would be a low quality source of high resolution information. The standards for high and low resolution will depend on the LOE and type of data being considered.

These three types of metadata represent only a few of the many possible factors that can be considered when contextualizing evidence and many additional factors may be considered. Other recent studies, for example, have included assessments of the strength of association between the measured data and the effect of interest, study design and execution, the degree to which the data has been reviewed, and evidence soundness, applicability, utility, clarity, completeness, uncertainty, and variability, etc. (Menzie et al. 1996; Linkov et al. 2011).

2.2 What is Weight of Evidence?

Weight of evidence (WOE) is a collection of qualitative and quantitative approaches used to interpret evidence in the context of its metadata. It is useful for synthesizing data from multiple sources to make a decision. WOE approaches can be used within an LOE to combine different pieces of evidence into a representative summary of that LOE's support for or against a hypothesis and the strength of that support. WOE approaches can also be used across LOEs (as demonstrated in our case study) to summarize the total support of diverse evidence for or against a hypothesis. Thus, WOE approaches help practitioners determining the extent to which each piece of evidence or LOE should influence the decision at hand. WOE analysis may also provide insights on the order in which LOEs should be considered. With respect to sediment management, these LOEs might include historic site usage data, analysis of benthic organisms for uptake of possible contaminants, computational modeling to evaluate contaminant mobility, etc. (for a more complete list, see Bridges et al. 2005; International Navigation Association 2006; Magar et al. 2009; Steevens 2013).

WOE approaches vary on a continuum from qualitative to quantitative, and on the type, amount, and complexity of information available (Table 1;

Smith et al. 2002; Weed 2005; Linkov and Satterstrom 2006; Linkov et al. 2009; Rhomberg et al. 2013). For example, Best Professional Judgment is a more qualitative approach, which entails listing different LOEs and holistically assessing the relevance of each one for the decision. For instance, an environmental engineer who considers how to remediate sediments at a port may initially place high importance on historical data describing the types of materials that have been transported through this site and modeling their mobility but may place little weight on benthic uptake information. In this case, if the historical and modeling evidence suggest that the material is toxic and mobile, Best Professional Judgment may suggest that consideration of benthic uptake is now also highly relevant to deciding how to handle the material. More quantitative approaches may attempt to numerically estimate the extent to which an LOE would inform a decision. For instance, a scoring rubric can be used to consistently convert judgments to scores that are then used to weight how much each LOE is considered relative to other LOEs when making a decision (Linkov et al. 2011).

Table 1. Comparison and summaries of some WOE approaches (see Linkov et al. 2009 for details).

WOE Approach	Description
Listing Evidence	Presentation of individual lines of evidence without attempt at integration
Best Professional Judgment	Qualitative integration of multiple lines of evidence
Causal Criteria	A criteria-based methodology for determining cause and effect relationships
Logic	Standardized evaluation of individual lines of evidence based on qualitative logic models
Scoring	Quantitative integration of multiple lines of evidence using simple weighting or ranking
Indexing	Integration of lines of evidence into a single measure based on empirical models
Quantification	Integrated assessment using statistical methods and/or decision analysis

WOE was initially introduced as an application of Bayesian analysis (Good 1960; Schultz and Borrowman 2011). Under this formulation, the decision maker has prior beliefs about a hypothesis that are updated to form posterior beliefs after new evidence is considered. The natural logarithm of the ratio of prior odds to posterior odds (called the Bayes factor) quantifies

the WOE. Various other quantitative and qualitative applications of WOE have been developed from those roots (Menzie et al. 1996; Weed 2005; Hope and Clarkson 2014; Linkov et al. 2009, 2015). The methods discussed here represent simplified quantitative approaches, but generally align with the intent of Bayesian logic as long as past and current evidence are considered alongside each other in developing the total WOE scores. The choice of a WOE method should reflect the decision needs, stakeholder interests, and data availability. Approaches generally differ in terms of degree of quantification, ability to produce results that are easily and consistency interpretable, ease of use, level of discrimination provided between alternatives, breadth of applicability, and transparency (Burton et al. 2002a).

Other government agencies have also been increasingly turning to WOE approaches to synthesize complex information for robust and transparent decision making. For example, the European Food Safety Authority (EFSA) is holistically embracing WOE in what is one of the most sophisticated governmental implementations of the approach. The agency has just released comprehensive guidance on the broad use of WOE in the types of scientific assessments under their jurisdiction. The public review draft of their guidance addresses the use of several quantitative and qualitative WOE methods, from simply listing evidence through formal decision analysis and statistical approaches, and includes case studies documenting their application (EFSA 2017). In another example, the international WOE guidance of the United Nations' (UN) Globally Harmonized System of Classification and Labelling of Chemicals, also provides a set of criteria for implementing WOE when evaluating chemical health hazards (UN 2011).

Within the US, the USEPA and its partner agencies use and recommend the use of WOE extensively. For example, the USEPA, US-NOAA National Marine Fisheries Service (NMFS), and US Fish and Wildlife Service (FWS) see WOE approaches as beneficial in pesticide consultations under the Endangered Species Act regarding potential likely adverse effects of chemicals on animal life (Hartl 2015; Hecht 2015). The USEPA also recently published a new, approx. 100-page report on the use of WOE in ecological risk assessment that provides guidance on using WOE through a standard framework of assembling evidence, weighting evidence, and interpreting the weight of the body of evidence (Suter 2016). In another example, their current guidelines for carcinogenic risk assessment, WOE is

used to combine human test data, animal test data, and other supporting evidence to characterize an agency's potential human carcinogenicity (USEPA 2005). The USEPA's endocrine disruption screening program also uses WOE to evaluate the results of first-tier risk screening to evaluate the need for further testing. Here, the agency outlines ways to use the results of different types of assays to evaluate risk pathways, and includes specification of information quality guidelines (USEPA 2011).

The US Occupational Safety and Health Administration (USOSHA) also recently issued guidance for public review on how to perform data evaluation for WOE determination, which are consistent with the UN chemical classification and labeling guidance. In their proposed guidance, tailored to assessing health hazards of chemicals, the use of expert interpretations of evidence in a WOE context is encouraged because it "provides a systematic way to evaluate a group of health effects studies that vary in quality and provide conflicting information," makes use of all available information regardless of type, and allows pooling of several less-conclusive studies to arrive at more-conclusive results (USOSHA 2016). These and other implementations show the importance of WOE to US and international government agencies for taking a comprehensive yet nuanced approach to transparently integrating diverse lines of evidence to reach broader conclusions, with transparent consideration of uncertainty.

2.3 Implementing a weight-of-evidence analysis

In general, some form of the following six steps will need to be implemented during a WOE analysis (adapted from Linkov et al. 2011). Depending on the decision context and the WOE approach used (Table 1), these steps may be implemented quickly, mentally, and qualitatively, or formally, quantitatively, and with deliberation:

1. Formulate decision objective
   In a WOE analysis, the objective of the current decision should be stated unambiguously. Vaguely stated objectives make it needlessly difficult to interpret and evaluate evidence in the context of the decision. A well-stated objective helps WOE implementers to more effectively combine information across dissimilar LOEs and identify LOE metadata that is relevant for the decision.
2. Identify alternatives and formulate hypotheses
   The decision under consideration involves a choice between alternatives. This might be a choice between management options (e.g.,

dredging at site A or B), courses of action (e.g., whether or not to remediate), or classification of an object (e.g., level of contamination in a soil sample), etc. For a comprehensive decision, all alternatives being considered should be listed prior to evaluation. Each alternative can be represented by a hypothesis asserting that that alternative makes a suitable choice for the decision. The evidence is then consulted (in step 5) to infer the strength of support for or against each hypothesis and the alternative it represents. For example, one hypothesis might assert that "the open water placement site is suitable for use in this dredging project." After consulting the evidence, the decision makers might find either strong or weak support for or against that hypothesis. The hypothesis with the strongest total support identifies the best alternative.

3. Structure decision objectives/explanatory factors through LOEs
   A hierarchical structure is a common way to decompose a general problem into more specific components. The main objective may be decomposable into sub-objectives or potential explanatory factors relating observed effects to potential causes. Ultimately, the lowest-level factors should be able to be linked to data that can be estimated or collected to evaluate the alternatives. This might include the myriad types of data collected during a complex ecological risk assessment (e.g., Hope and Clarkson 2014). The LOEs typically correspond to the lowest-level factors, or to multiple types of evidence within a factor, because these each offer different types of support for or against a hypothesis.
4. Structure metadata
   Metadata surrounding the evidence should influence how strongly we believe (or place weight in) the claims evidence makes for or against a hypothesis. For example, LOEs measuring grain size, contaminant concentration, and placement costs, might benefit from consideration of metadata about instrument error, spatial and temporal resolution, or the precision of estimates used in the analysis. The most relevant metadata for interpreting evidence in the context of the decision at hand needs to be listed.
5. Gather data to evaluate the hypotheses (assess the alternatives)
   Data are used to assess the strength of support for accepting or rejecting each hypothesis about an alternative being suitable for selection in the decision. This assessment involves considering both what the evidence states about the hypothesis and what the metadata suggests about the perceived strength of those statements. This might

involve two dimensions of hypothesis and metadata scores, or listed evaluation considerations in a table summarizing all available information for each alternative.

6. Synthesize Information (and Sensitivity Analysis)
Once all the evidence and metadata has been compiled, it needs to be considered together to reach conclusions. Each hypothesis can be evaluated independently, and the evaluation can be either qualitative or quantitative. If done quantitatively, the individual hypothesis and metadata scores can be aggregated into a total LOE score for each lowest-level factor or type of evidence, and then aggregated across decision objectives or factors to provide a total WOE score for each hypothesis about alternative suitability. If the decision sub-objectives/factors (in step 3) or metadata properties (in step 4) are not of equal importance for the decision, trade-offs or priority weights can be applied during the aggregation. If aggregation does not seem appropriate for a particular decision, Menzie et al. (1996) suggest plotting the total LOE scores for visual inspection and to identify areas of relative agreement or differences across evidence types. After arriving at a conclusion, sensitivity analyses can be conducted to understand how the weight of evidence might differ if other factors were considered in the analysis or if the factors were considered with different levels of importance.

2.4 Brief summary of four WOE applications

The following paragraphs briefly summarize four WOE applications that show the diversity of WOE approaches used for environmental studies; how the objectives, criteria, and alternatives were structured in each case; the range of LOEs and metadata that were considered; and how WOE was able to provide helpful decision support in different contexts.

Khosrovyan et al. (2015) consider two WOE approaches, both of which evaluate sediment quality and the toxicity risk posed to aquatic biota. The first uses a decision tree to guide the collection and interpretation of multiple and increasingly detailed pieces of evidence. This helps users achieve an appropriate level of detail for each decision. For example, the result of assessing a LOE could lead to either termination of the risk characterization or a suggestion to pursue additional LOEs for the evaluation. The second approach attempts to statistically attribute toxicity response to environmental variables. Data from multiple lines of evidence are integrated by principal component analysis and components are

subsequently correlated with environmental variables. When applied to the same dataset, the two methods resulted in relatively consistent assessments of sediment toxicity risk. In this case, the choice to integrate WOE concepts into decision making seems to have been more important than the choice of which method to use.

Lowell et al. (2000) use WOE to address the challenge of "distinguishing among the cumulative impacts of multiple stressors" in a northern riverine system. Their approach identifies the link between observed ecosystem conditions and the cause(s) of impairments, informing management action. From the literature, the authors establish a set of criteria for assessing potential causes of river impairment. Called causal criteria, this set facilitates evaluation of the available information to determine primary stressor/effect(s). The criteria they used include spatial and temporal correlation of stressor and effect; plausibility of stressor/effect link; existence of experimental verification; strength, specificity, and consistency of evidence; and coherence with existing knowledge. In their study of the complex riverine system's dynamics, satisfying the WOE criteria established sufficient causality to support consideration of multiple concurrent effects. The ability of their framework to combine evidence across multiple studies and rivers was key to gaining insights into the effects of stressors so that they could be effectively managed.

Smith et al. (2002) note that considering multiple LOEs pertaining to sediment quality is challenging and that combining them into a single measure is valuable for comparing and ranking management sites. To this end, they implement a statistical WOE approach that estimates the probability of impairment for a site based on probability distributions for each LOE and conditions at a reference sites. The authors consider a sediment quality triad (Chapman 1996) of three LOEs, including toxicity, biological field, and chemistry. These are weighted numerically to define "how much an observed feature in the data adds to or subtracts from the evidence of impact." The probabilities of impairment are combined across LOEs to generate an overall probability of site impairment. The authors apply this WOE approach to a contaminated ship repair site in Lake Huron, which has since been remediated.

Suter & Cormier (2011) outline a generic approach for WOE inference in which hypotheses are supported by strong evidence or revised if evidence is weak. Their process consists of weighing each piece of evidence in a LOE

and weighting each LOE as a whole to produce some indication of the strength or weakness of each candidate hypothesis for an observed state. In their demonstration, they strive to find the cause of a decline in population abundance of the San Joaquin kit fox. Each piece of evidence is weighted based on whether it strengthens or weakens support for a hypothesized cause for the decline. For example, a piece of observed data may support a hypothesis that habitat disturbance is the cause of decline. LOEs are then weighted using criteria of “consistency” and “reasonableness of explanation.” This type of weighting is attuned to the implications of evidence, but does not consider other factors, such as the quality of evidence, for example. It supports the comparison and consideration of alternative hypotheses for explaining an observed phenomenon — which, the authors note, is a less common approach for conducting WOE assessments but may be more reliable.

3 Case Study: Introduction

The following case study walks through the steps of applying one WOE approach to a dredged material management decision. This example focuses on part of a dredging project for New Haven Harbor, CT, the busiest port on Long Island Sound and one of the busiest ports in New England. The harbor is operating at reduced efficiency since shoaling allows large vessel access to the inner harbor only at high tide. After analysis, project managers deemed it necessary to dredge to the full federally authorized depth in order to maintain local welfare and economic efficiency. In this case study, information from New Haven Harbor is used to propose and evaluate hypotheses about what to do with sediment generated by this dredging activity. Four hypotheses are evaluated, proposing that four different placement areas (that passed an initial screening) are suitable for use with the New Haven Harbor dredging project. Figure 1 shows the New Haven Harbor area with some sediment-testing locations identified. Figure 2 summarized how the general WOE concept is applied. This shows all eight alternatives initially considered, though only four moved forward to the full WOE analysis after passing the initial feasibility screening. A mix of qualitative and quantitative information summarizing relevant evidence and metadata is used to generate total WOE scores for each hypothesis.

Figure 1. New Haven Harbor, CT, with shoal removal areas and sampling locations (Martin 2013).*

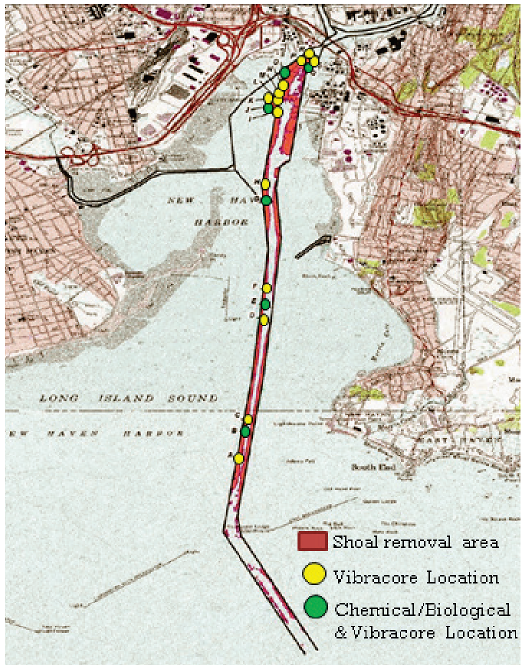
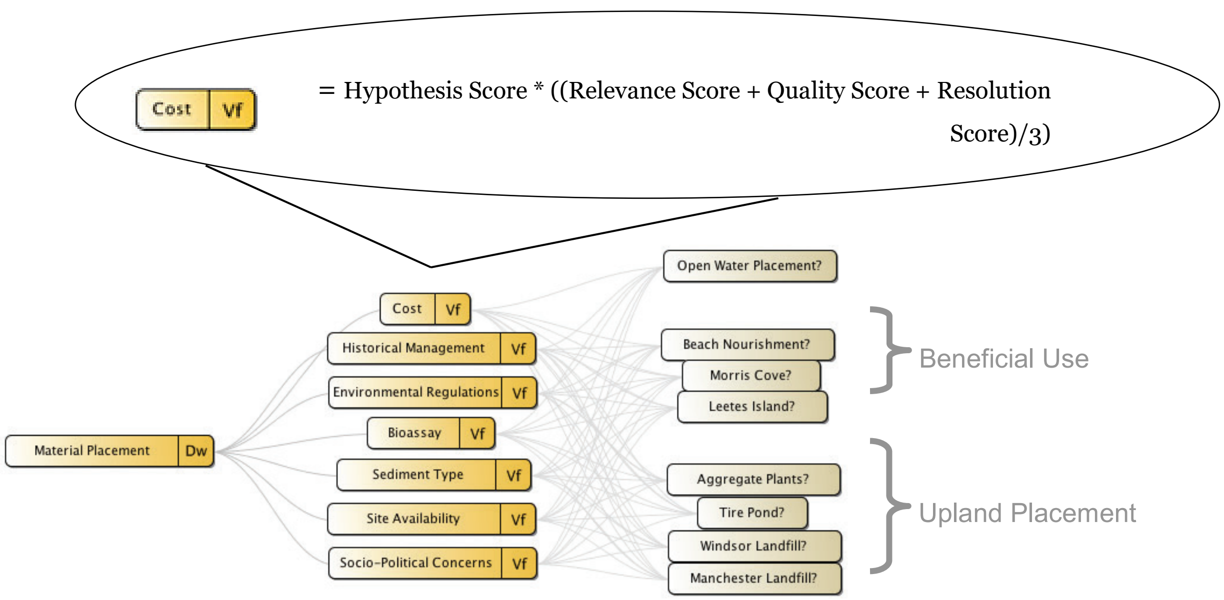


Figure 2. The material placement WOE model, showing each placement alternative being evaluated against multiple criteria based on its degree of support for or against a hypothesis about a site's suitability.



* Martin, Craig. 21 February 2013. *New Haven CT Federal Navigation Project Long Island Sound Regional Dredge Team*. Presentation given for the New England District U.S. Army Corps of Engineers.

In addition to the hypothesis score, the evidence is also evaluated with respect to metadata about its relevance, quality, and resolution for the decision. The hypothesis and metadata scores for alternatives not ruled out as infeasible are combined into a total WOE score for each alternative.

3.1 Case Study: Dredged material placement alternatives

An initial screening was conducted by the USACE district to make sure that evaluated alternatives met minimum feasibility standards. Site availability constraints were found to make placement infeasible at a proposed tire pond, the aggregate plants, and the Windsor landfill sites. These sites were removed from further analysis. The beach nourishment alternative was also removed due to material incompatibility, since the dredged material contained mostly silts and clays. The four remaining alternatives are analyzed in a WOE process. These sites included:

1. Manchester Landfill
   The Manchester landfill site is large enough to receive the projected quantity of material. However, how it would fill roughly 75% of the remaining available space at the landfill and the extent of the already-contracted area was unclear (Martin 2013; thus, the authors give a med-low metadata quality score to data suggesting that the site is available). Landfill managers also had concerns that the quantity and contents of material would face substantial resistance from environmental stakeholders in the city of Hartford and surrounding towns (thus, the authors give only a medium relevance score to the bioassay data, since other environmental concerns would likely dominate). This alternative was projected to carry a high cost on the order of $120 million, representing an approximately 1390% overall cost increase from the lowest cost alternative after an extensive cost analysis (Martin 2013). Significant infrastructure would also have to be constructed to dewater the dredged material.
2. Central Long Island Sound (CLIS) Site
   Open water placement of material at the CLIS site, approximately 8 kilometers (5 miles) from the dredging location, was found to be the lowest cost option at $8.6 million U.S. dollars (Martin 2013). The primary cost involved with this alternative is transporting material via scows to the placement site. This alternative was also known to be feasible since it had already been undertaken eight times (from historical

dredged material placement records; Martin 2013). Based on its prior use, it was unlikely that placement would generate substantial new socio-political concerns. Placement operations would be postponed during the shellfish spawning season from June to September.

3. Morris Cove
   Morris Cove contains a borrow pit that was dug for the construction of highway I-95 in the 1950s. The NOAA designates this region as Essential Fish Habitat and it currently has decreased productivity due to low oxygen levels during summers. Placement of new material would improve water quality and likely restore shellfish beds (USACE EA 2013). Sediment compatibility for this site would be highly relevant and was found to be compatible based on (high resolution) surveys conducted in 2011. About 75% of the approximately 460,000 total cubic meters (600,000 cubic yards) of material could be placed at Morris Cove (Martin 2013). However, a secondary site would also be required (medium availability). The scow would also have to be light loaded in order to reach the cove for material placement, thus leading to decreased dredging productivity and longer timelines. The overall cost is moderate at $10.8 million U.S. dollars (Martin 2013).
4. Leetes Island
   Leetes Island is located approximately 29 kilometers (18 miles) from the dredging site. Material placed here would raise the marsh surface and support restoration. Sediment compatibility would be highly relevant and was found compatible. The additional travel distance to the site would increase emissions due to the 1,400 required truck trips. In addition, the island only has the capacity to house 3.5% of the total dredged material (low site availability). A secondary placement location for the other 96.5% of the material would also be required, here assumed to be the CLIS site. Total projected cost was moderately high at approximately $13.8 million.

3.2 Case Study: LOES for dredged material placement

Seven LOEs were considered in this analysis to illustrate the basics of a WOE approach. For simplicity the authors considered only one piece of evidence within each type of LOE. Hypotheses were proposed suggesting that each potential placement site was suitable for use in the New Haven Harbor dredging project, and each LOE was rated based on whether it highly supported, supported, was indeterminate with respect to, opposed, or highly opposed each hypothesis. These ratings were converted into hypothesis scores using the scoring rubric in Table 1. LOE metadata were

also reviewed for relevance, quality, and resolution, and rated as being of either high, med-high, med, med-low, or low strength based on author judgment of the evidence in its full context. These ratings were converted to scores using the rubric in Table 2. While many of the scores differ between alternatives, some are uniform because similar techniques were applied to all sites considered; these details are discussed in the text below. The LOEs considered in this case study include:

1. Sediment type analysis
   Nineteen vibracore samples were taken in the navigation channel at points around the inner and outer harbor (Figure 1). Samples from each of these were analyzed for the type of material contained in the sediment as a proportion of total mass. Material in each sample are classified as Silt/Clay, Fine Sand, Medium Sand, Course Sand, Very Course Sand, or Gravel. Percentage make-ups for each type were also provided and sediment type compatibility is considered for each of the alternative placement sites.
2. Contamination analysis through bioassay
   An in vivo bioassay analysis was performed for six samples using clams (*Macoma nasuta*) and worms (*Nereis virens*) with sediment from the vibracore samples in order to evaluate the risk of potential contaminants to organisms in each placement area. Three separate tests were undertaken with timelines ranging from 10-28 days (U.S. Army Corps of Engineers EA, 2013). PCBs and DDT were found in the samples at relatively low concentrations.
3. Cost
   Dredged material placement measures were subject to extensive cost analysis, incorporating all known aspects of the placement procedures for each site. Percentage cost increases between alternatives were also quantified.
4. Environmental effects
   The *federal standard* for costs and environmental impacts on a dredging project is required to be met. The dredge type used for the proposed placement must comply with section 176(c) of the Clean Air Act (CAA; U.S. Army Corps of Engineers EA 2013). Water quality modeling was also conducted to ensure that existing criteria for water quality standards would be met during the project. (A complete list of statute compliance is included in the site-wide EA.) Site managers of potential placement areas were also contacted to discuss potential environmental concerns.

5. Historical management of dredged material
   Records indicate that the CLIS site was previously used eight times as a dredged material disposal site. CLIS is located roughly 8 kilometers (5 miles) from the dredge site. An extensive Environmental Assessment (EA) was conducted with a finding of no significant impact for dredged material disposal at the CLIS site. The other considered placement sites have not been historically used.
6. Site availability
   Managers of potential placement sites were contacted to determine availability for the quantity of material to be placed in accordance with proposed timelines for the project.
7. Socio-political concerns
   Possible sites for material placement were analyzed for potential socio-political concerns. Site managers and operators would be expected to refuse or be less likely to accept material if they believe the public would disapprove of such actions.

3.3 Case Study: WOE model application

There are many ways to implement a WOE analysis. In this case study, the authors use a mix of Best Professional Judgment (BPJ) and the Multi-Criteria Decision Analysis (MCDA) approach outlined by Linkov et al. (2011). BPJ helps summarize and interpret available knowledge where quantitative data is sparse or poorly matched to the decision factors of interest; MCDA (Belton & Stewart 2002) helps decision makers quantify the extent to which each line of evidence contributes to a conclusion. In this example, the authors use BPJ to first evaluate the LOEs on cardinal scales (e.g., high/medium/low), and then generate MCDA scores that quantify relative contributions to the decision.

The authors use BPJ to qualitatively assess LOE support for a hypothesis ranging from “Highly Opposes” to “Highly Supports.” These responses are quantified into *Hypothesis Scores* ranging from -1 to 1. Similarly, metadata for evidence quality, relevance, and resolution are first assessed qualitatively on a range from “Provides No Information” to “Provides Perfect Information.” These responses are then quantified into *Metadata Scores* between 0 and 1. Tables 2 and 3 show the rubrics used to translate the qualitative value descriptors into quantitative scores.

Table 2. Hypothesis scoring rubric transforms BPJ qualitative assessments into quantitative scores representing evidence support. When integrated with the metadata scores, these will provide a total score for each LOE and a total WOE score for each alternative.

Description	Hypothesis Score
HIGHLY SUPPORTS	1.0
SUPPORTS	0.5
INDETERMINATE	0.0
OPPOSES	-0.5
HIGHLY OPPOSES	-1.0

Table 3. Metadata scoring rubric transforms BPJ qualitative assessments into quantitative scores representing LOE weight. When integrated with the hypothesis scores, these will provide a total score for each LOE and a total WOE score for each alternative. In the New Haven Harbor example, metadata scores are assessed in terms of data quality, relevance, and resolution. Note that, while included as the 0 and 1 logical anchors of the metadata rubric, discovering evidence that is entirely perfect or useless is uncommon in most practical settings and, if found, would not require WOE to interpret. Pragmatic metadata scores range from a lower bound of 0.1 to an upper bounds of 0.9, with a middle score of 0.5.

Description	Metadata Score
(PERFECT)	*(1.0)*
HIGH	0.9
MED-HIGH	0.7
MED	0.5
MED-LOW	0.3
LOW	0.1
(NONE)	*(0.0)*

MCDA-based equations are used for quantifying the total LOE and WOE scores used in this case study. Note that these equations assume equal weighting among the different metadata properties and LOE criteria. (To differentiate the relative importance of the criteria, a weighting multiplier can be used within the summations instead of dividing by the number of items). Note that any BPJ response of “N/A” would be omitted from the averages in these equations:

$$Total\ LOE\ Score\ for\ an\ alternative = \frac{\sum Metadata\ Scores}{\#\ Metadat\ Properties} * Hypothesis\ Score$$

$$Total\ WOE\ Score\ for\ an\ alternative = \frac{\sum Total\ LOE\ Scores}{\#\ LOE}$$

The appendix provides data tables showing all qualitative assessments, the quantified hypothesis and metadata scores, and the total LOE and WOE scores for each alternative.

3.4 Case Study: Results

The CLIS placement option scores highest in the WOE analysis (nearly twice as high as the next alternative), suggesting it as the most suitable placement site. The Morris Cove option scores next, followed by the Leetes Island and the Manchester Landfill (see the tables in the Appendix for details). It is worth noting that these case study results agree with the decision that project managers ultimately made to select the CLIS site for the actual dredged material placement.

WOE for dredged material placement can help decision makers understand the feasibility of possible alternatives. The New Haven Harbor case study provides one example of how multiple LOEs can be combined in a WOE analysis, where data are interpreted in context to help better discriminate between what evidence says and how much emphasis we should place on it in the decision. Bridging the data-to-decision gap transparently and quantitatively makes the analysis reproducible, enables future sensitivity and scenario analysis, and increases the actual and perceived rigor and robustness of the results. To add additional nuance to this analysis, non-equal weighting could to be used to reflect the preferences of the stakeholder community or the decision makers. The analysis could also be implemented via a Bayesian WOE model (Good 1960), which, while more difficult to implement, could provide better granularity in the results.

4 General Conclusions

This special report introduces the basic concepts of WOE analysis and shows, via brief summaries and a case study, how it can be applied to USACE Civil Works projects. By combining various LOEs that are weighted based on their relative importance, users can form a transparent and comprehensive decision-making framework that aggregates diverse information in context. WOE offers a unique contribution for decision makers through its inclusion of metadata concerns (e.g., quality, relevance, and resolution in the case study) and its ability to explicitly consider the credibility of the data used in the context of the decision at hand.

As science and society advance, additional calls are being made for decision transparency. WOE approaches can reveal the logic that decision makers use to reach their ultimate conclusions, which builds additional stakeholder trust. WOE approaches present a flexible but rigorous set of qualitative and quantitative methods that can offer clarity on how challenging dilemmas were addressed. WOE approaches have been widely used across many domains, have solid scientific underpinnings, and are recommended for broader adoption throughout USACE.

4.1 Further reading

For additional applications of WOE in case studies, see the classification of polluted sediments by Benedetti et al. (2011); the application of WOE to assess sediment contamination in Burton et al. (2002b); the evaluation of diagnosing aquatic system impairment by Kapo and Burton (2006); the evaluation of sediment contamination and remediation strategies by Linkov et al. (2011); and the additional study of probability for sediment contamination in the Great Lakes by Smith et al. (2002). For additional review and discussion of WOE in the context of environmental risks and ecosystem impairment, see Menzie et al. (1996); Lowell et al. (2000); Burton et al. (2002a); Hull and Swanson (2006); Suter and Cormier (2011); and Hope et al. (2014).

References

Belton, V., Stewart, T. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Springer Science and Business Media.

Benedetti, M., F. Ciaprini, F. Piva, F. Onorati, D. Fattorini, A. Notti, A. Ausili, F. Regoli. 2011. A multidisciplinary weight of evidence approach for classifying polluted sediments: Integrating sediment chemistry, bioavailability, biomarkers response, and bioassays. *Environment International* 38: 17-28.

Bridges, T. S., W. J. Berry, S. Della Sala, P. B. Dorn, S. J. Ells, T. H. Gries, D. S. Ireland, E. M. Maher, C. A. Menzie, L. M. Porebski, J. Stronkhorst. 2005. A framework for assessing and managing risks from contaminated sediments. In *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments,* ed. R. J. Wenning, G. E. Batley, C. G. Ingersoll and D. W. Moore, 227-266. Pensacola, FL: SETAC Press.

Burton, G. A., P. M. Chapman, and E. P. Smith. 2002a. Weight-of-evidence approaches for assessing ecosystem impairment. *Human and Ecological Risk Assessment* 8(7): 1657-1673.

Burton, G. A., G. E. Batley, P. M. Chapman, V. E. Forbes, E. P. Smith, T. Reynoldson, C. E. Schlekat, P. J. den Besten, A. J. Bailer, A. S. Green and R. L. Dwyer. 2002b. A weight-of-evidence framework for assessing sediment (or other) contamination: Improving certainty in the decision-making process. *Human and Ecological Risk Assessment 8*(7): 1675-1696.

Chapman, P. M. 1996. Presentation and interpretation of Sediment Quality Triad data. *Ecotoxicology* 5: 327-339.

Good, I. J. 1960. "Weight of evidence, corroboration, explanatory power, information and the utility of experiments." *Journal of the Royal Statistical Society* 22: 319-331.

Hope, B. K., and J. R. Clarkson. 2014. A strategy for using weight-of-evidence methods in ecological risk assessments. *Human and Ecological Risk Assessment: An International Journal 20*(2): 290-315.

Hull, R. N., and S. Swanson. 2006. Sequential analysis of lines of evidence—an advanced weight-of-evidence approach for ecological risk assessment. *Integrated Environmental Assessment and Management 2*(4): 302-311.

International Navigation Association (PIANC). 2006. Biological assessment guidance for dredged material. EnviCom report of WG 8 - January 2006 issue. Brussels, Belgium.

Kapo, K. E., and G. A. Burton. 2006. Geographic information systems-based, weights-of-evidence approach for diagnosing aquatic ecosystem impairment. *Environmental Toxicology and Chemistry* 25(8): 2237-2249.

Khosrovyan, A., A. Rodríguez-Romero, M. A. Ramos, T. A. DelValls, and I. Riba. 2015. Comparative analysis of two weight-of-evidence methodologies for integrated sediment quality assessment. *Chemosphere* 120: 138-144.

Linkov, I., and F. K. Satterstrom. 2006. Weight of evidence: What is the state of the science? *Risk Analysis* 23(3): 573-575.

Linkov, I., D. Loney, S. Cormier, F. K. Satterstrom, and Bridges, T. 2009. Weight-of-evidence evaluation in environmental assessment: Review of qualitative and quantitative approaches. *Science of the Total Environment* 407: 5199-5205.

Linkov, I., P. Welle, A. Tkachuk, L. Canis, J. B. Kim, T. Bridges. 2011. Use of multicriteria decision analysis to support weight of evidence evaluation. *Risk Analysis* 31(6): 1211-1225.

Linkov, I., O. Massey, J. Keisler, I. Rusyn, and T. Hartung. 2015. From" weight of evidence" to quantitative data integration using multicriteria decision analysis and Bayesian methods. *Altex* 32(1): 3.

Lowell, R. B., J. M. Culp, and M. G. Dubé. 2000. A weight-of-evidence approach for northern river risk assessment: Integrating the effects of multiple stressors. *Environmental Toxicology and Chemistry* 19(4): 1182-1190.

Magar, V. S., D. B. Chadwick, T. S. Bridges, P. C. Fuchsman, J. M. Conder, T. J. Dekkar, J. A. Steevens, K. E. Gustavson, and M. A. Mills. 2009. *Technical guide: monitored natural recovery at contaminated sediment sites*. ESTCP Project ER-0622. Alexandria, VA: Environmental Security Technology Certification Program.

Menzie, C., M. H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risks. *Human and Ecological Risk Assessment* 2(2): 277-304.

Rhomberg, L. R., J. E. Goodman, L. A. Bailey, R. L. Prueitt, N. B. Beck, C. Bevan, M. Honeycutt, N. E. Kaminski, G. Paoli, L. H. Pottenger, R. W. Scherer, K. C. Wise and R. A. Becker. 2013. A survey of frameworks for best practices in weight-of-evidence analyses. *Critical Reviews in Toxicology* 43(9): 753-784.

Schultz, M. T. and T. D. Borrowman. 2011. *An Introduction to using Bayesian networks to model dredging decisions*. Vicksburg, MS: U.S. Army Engineer Research and Development Center. Available at: http://acwc.sdp.sirsi.net/client/en_US/search/asset/1005525

Smith, E. P., I. Lipkovich, K. Ye. 2002. Weight-of-evidence: Quantitative estimation of probability of impairment for individual and multiple lines of evidence. *Human and Ecological Risk Assessment* 8(7): 1585-1596.

Steevens, J. A. July 2013. *Overview of Dredged Material Testing and Evaluation*. Presentation given as part of the Dredging Operations Technical Support program Summer Webinar Series, Vicksburg, MS. Available at: http://el.erdc.usace.army.mil/dots/training.html

Suter, G. W., and S. M. Cormier. 2011. Why and how to combine evidence in environmental assessments: weighing evidence and building cases. *Science of the Total Environment* 409(8): 1406-1417.

U.S. Army Corps of Engineers New England District. (29 January 2013). *Public notice: 30 day public notice maintenance of the New Haven Harbor Federal Navigation Project, New Haven, Connecticut*. Available at: http://www.nae.usace.army.mil/portals/74/docs/Navigation/NewHaven29Jan13.pdf

______. June 2013. *Environmental assessment finding of no significant impact, Clean Water Act section 404(b)(1) evaluation*. New Haven Harbor: New Haven, CT.

USEPA and USACE. 1991. *Evaluation of dredged material proposed for ocean disposal*. Report # EPA 503/8-91/001. Washington, DC: USEPA.

_______. 1998. Evaluation of dredged material proposed for discharge in waters of the U.S. – Testing manual. Report # EPA-823-B-98-004. Washington, DC: USEPA.

Weed, D. L. 2005. Weight of evidence: A review of concept and methods. *Risk Analysis* 25(6): 1545-1557.

Appendix: Case Study Data, Calculations, and Results

MANCHESTER LANDFILL

Lines of Evidence	Metadata Scores			Average Metadata Score	Hypothesis Score	Total LOE Score
	Relevance	Quality	Resolution			
Cost	HIGH (0.9): *Budget constraints were required to be met*	HIGH (0.9): *Extensive holistic cost analysis*	MED-HIGH (0.7): *4 Significant figures used*	0.83	HIGHLY OPPOSES (-1.0): *Extremely high cost- $119M= 1390% increase*	-0.83
Historical Management	MED-LOW (0.3): *Not a required or limiting factor but is informative*	LOW (0.1): *Records limited*	N/A	0.2	NEUTRAL (0): *Landfill placement rarely used if other measures are deemed feasible*	0
Environmental Effects	HIGH (0.9): *Federal regulations were required to be met for all project aspects*	LOW (0.1): *Environmental effects for placement are vague*	N/A	0.5	HIGHLY OPPOSES (-1.0): *EPA does not recommend Huge emissions increase*	-0.5
Bioassay	MED (0.5): *Landfill owners would likely only be concerned with high levels of contamination*	HIGH (0.9): *3 different in vivo tests*	MED-LOW (0.3): *Sample size of 6*	0.57	NEUTRAL (0.0): *Low levels of PCBs and DDTs detected*	0
Sediment Type	LOW (0.1): *Sediment compatibility is not a concern for landfill owners*	MED (0.5): *Vibracore considered 2nd tier*	HIGH (0.9): *Sample size of 19*	0.5	HIGHLY SUPPORTS (1.0): *Relatively any sediment type would be compatible*	0.5
Site Availability	HIGH (0.9): *Key requirement for placement*	MED-LOW (0.3): *Quantity of already contracted capacity unknown*	N/A	0.6	HIGHLY SUPPORTS (1.0): *Available for the full quantity of material*	0.6
Socio-Political	MED-HIGH (0.7): *Public unrest could make an alternative less feasible*	MED-LOW (0.3): *Speculation, no known polling occurred*	N/A	0.5	HIGHLY OPPOSES (-1.0): *Predicted to be a large upset to local stakeholders*	-0.5
					TOTAL WOE SCORE:	**-0.10**

Reminder: Total WOE Score for an alternative $= \sum\left(\frac{\sum Metadata\ Scores}{\#\ Metadat\ Properties} * Hypothesis\ Score\right) / \#\ LOE$

CLIS SITE

Lines of Evidence	Metadata Scores			Average Metadata Score	Hypothesis Score	Total LOE Score
	Relevance	Quality	Resolution			
Cost	HIGH (0.9): *Budget constraints were required to be met*	HIGH (0.9): *Extensive holistic cost analysis*	MED-LOW (0.3): *Only 2 significant figures in cost estimate*	0.7	HIGHLY SUPPORTS (1.0): *Lowest Cost Option*	0.7
Historical Management	MED-LOW (0.3): *Not a required nor limiting factor but is informative*	HIGH (0.9): *Extensive records of previous dredged material placement were utilized*	HIGH (0.9): *Executed 8 times already*	0.7	HIGHLY SUPPORTS (1.0): *Previously executed successfully*	0.7
Environmental Effects	HIGH (0.9): *Federal regulations were required to be met for all project aspects*	HIGH (0.9): *An EA was conducted fully analyzing a breadth of possible effects*	N/A	0.9	SUPPORTS (0.5): *Recommended by EPA Finding of no significant impact*	0.45
Bioassay	HIGH (0.9): *Contaminants could unbalance the ecosystems*	HIGH (0.9): *3 different in vivo tests*	MED-LOW (0.3): *Sample size of 6*	0.7	NEUTRAL (0.0): *Some PCBs+ DDTs detected*	0
Sediment Type	HIGH (0.9): *Discrepancies in sediment could unbalance ecosystems*	MED (0.5): *Vibracore considered 2nd tier*	HIGH (0.9): *Sample size of 19*	0.8	HIGHLY SUPPORTS (1.0): *Compatible*	0.8
Site Availability	HIGH (0.9): *Key requirement for placement*	HIGH (0.9): *Stated in a NEPA impact statement*	N/A	0.9	HIGHLY SUPPORTS (1.0): *Fully available*	0.9
Socio-Political	MED-HIGH (0.7): *Public unrest could make an alternative less feasible*	MED-LOW (0.3): *Speculation, no known specific polling*	N/A	0.5	HIGHLY SUPPORTS (1.0): *Won't interfere with public activities; e.g., with shellfish*	0.5
					TOTAL WOE SCORE:	**0.58**

MORRIS COVE

Lines of Evidence	Metadata Scores			Average Metadata Score	Hypothesis Score	Total LOE Score
	Relevance	Quality	Resolution			
Cost	HIGH (0.9): *Budget constraints were required to be met*	HIGH (0.9): *Extensive holistic cost analysis*	MED (0.5): *3 significant figures used*	0.8	NEUTRAL (0.0): *$10.8M= 125% increase*	0.0
Historical Management	MED-LOW (0.3): *Historic use is not a required or limiting factor but is informative*	N/A	N/A	0.3	NEUTRAL (0.0): *Not previously used*	0.0
Environmental Effects	HIGH (0.9): *Federal regulations were required to be met for all project aspects*	MED (0.5): *Vague descriptions of environmental surveys for total impact*	N/A	0.7	HIGHLY SUPPORTS (1.0): *Supporting Essential Fish Habitat; Lower emissions*	0.7
Bioassay	HIGH (0.9): *Contaminants could destroy essential habitat*	HIGH (0.9): *3 different in vivo tests*	MED-LOW (0.3): *Sample size of 6*	0.7	NEUTRAL (0.0): *Low levels of PCBs and DDTs detected*	0.0
Sediment Type	HIGH (0.9): *Incompatible sediment could destroy essential habitat*	MED (0.5): *Vibracore considered 2nd tier*	HIGH (0.9): *Sample size of 19*	0.8	HIGHLY SUPPORTS (1.0): *Compatible*	0.8
Site Availability	HIGH (0.9): *Key requirement for placement*	MED (0.5): *Recent surveys conducted for site availability*	N/A	0.7	NEUTRAL (0.0): *Will accept only 75% of total material*	0.0
Socio-Political	MED-HIGH (0.7): *Public unrest could make an alternative less feasible*	MED-LOW (0.3): *Speculation, no known specific polling*	N/A	0.5	NEUTRAL (0.0): *No new infrastructure; Longer project timeline*	0.0
					TOTAL WOE SCORE:	**0.21**

LEETS ISLAND

Lines of Evidence	Metadata Scores			Average Metadata Score	Hypothesis Score	Total LOE Score
	Relevance	Quality	Resolution			
Cost	HIGH (0.9): *Budget constraints were required to be met*	HIGH (0.9): *Extensive holistic cost analysis*	MED (0.5): *3 significant figures in cost estimate*	0.8	OPPOSES (-0.5): *$13.8M= 160% cost increase*	-0.4
Historical Management	MED-LOW (0.3): *Not a required or limiting factor but is informative*	N/A	N/A	0.3	NEUTRAL (0.0): *Not previously used*	0.0
Environmental Effects	HIGH (0.9): *Federal regulations were required to be met for all project aspects*	LOW (0.1): *Sparse supporting information for impact is found*	N/A: *Sample size for surveys is unclear*	0.33	NEUTRAL (0.0): *Increases emissions to travel to site (1,400 truck loads 29 km away); Supports marsh restoration*	0.0
Bioassay	HIGH (0.9): *Contaminants could destroy recovering, critical habitat*	HIGH (0.9): *3 different in vivo tests*	MED-LOW (0.3): *Sample size of 6*	0.7	NEUTRAL (0.0): *Low levels of PCBs and DDTs detected*	0.0
Sediment Type	HIGH (0.9): *Incompatible sediment could destroy critical habitat*	MED (0.5): *Vibracore considered 2nd tier*	HIGH (0.9): *Sample size of 19*	0.8	SUPPORTS (0.5): *Compatible*	0.4
Site Availability	HIGH (0.9): *Key requirement for placement*	MED-LOW (0.3): *Information is from a vague source*	N/A	0.6	HIGHLY OPPOSES (-1.0): *Can only accept 3.5%, requires additional site*	-0.6
Socio-Political	MED-HIGH (0.7): *Public unrest could make an alternative less feasible*	MED-LOW (0.3): *Speculation, no known specific polling*	N/A	0.5	NEUTRAL (0.0): *Would support marsh restoration; Requires new infrastructure*	0.0
					TOTAL WOE SCORE:	**-0.09**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
March 2018	Final Special Report	**Jan 2013 - June 2017**

4. TITLE AND SUBTITLE

Weight-of-Evidence Concepts: Introduction and Application to Sediment Management

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Matthew E. Bates, Olivia C. Massey, Matthew D. Wood

5d. PROJECT NUMBER
406028

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Environmental Laboratory, US Army Engineer Research and Development Center
696 Virginia Road, Concord, MA 01742

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, MA 02139.

8. PERFORMING ORGANIZATION REPORT NUMBER

ERDC/EL SR-18-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Dredging Operations and Environmental Research (DOER) Program
US Army Engineer Research and Development Center (ERDC)
3909 Halls Ferry Road, Vicksburg, MS 39180-6199

10. SPONSOR/MONITOR'S ACRONYM(S)

DOER

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
SR 18540

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The U.S. Army Corps of Engineers (USACE) has a variety of information available to support Civil Works project planning and operations. While this information must be interpreted to inform conclusions, clear guidance is often not available to describe how best to compare and integrate different types of information. This special report introduces a Weight of Evidence (WOE) approach that USACE staff can use to interpret the many Lines of Evidence (LOEs) available in the information based on the conclusions that they support and how much weight they each should have in the decision, helping to bridge the data-to-decisions gap. A case study is presented applying WOE to evaluate potential sediment placement sites for dredged material from New Haven Harbor, CT, the busiest port on Long Island Sound and one of the busiest ports in New England. The analysis demonstrates how diverse pieces of information can be evaluated based on their quality, resolution, and relevance and combined in the context of the problem at hand to aid robust and trans-parent USACE Civil Works decision making.

15. SUBJECT TERMS

Civil works, Decision analysis, Dredge program, Information synthesis, Metadata, Transparent, Weight-of-Evidence

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**a. ABSTRACT**	**b. THIS PAGE**			Matthew E. Bates
Unlimited	Unlimited	Unlimited	SAR	35	**19b. TELEPHONE NUMBER** *(Include area code)* 978-318-8795